

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Relationship between time interval from primary surgery to the start of taxane- plus platinum-based chemotherapy and clinical outcome of patients with advanced epithelial ovarian cancer: results of a multicenter retrospective Italian study

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/36367> since

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

This is the author's final version of the contribution published as:

Berchialla P;Scarlinzi C;Snidero S;Gregori D. Comparing models for quantitative risk assessment: an application to the European Registry of foreign body injuries in children.. STATISTICAL METHODS IN MEDICAL RESEARCH. None pp: ---.

DOI: 10.1177/0962280213476167

The publisher's version is available at:

<http://smm.sagepub.com/cgi/doi/10.1177/0962280213476167>

When citing, please refer to the published version.

Link to this full text:

<http://hdl.handle.net/2318/139382>

Comparing models for quantitative risk assessment: an application to the European Registry of foreign body injuries in children

Paola Berchialla,¹ Cecilia Scarinzi,² Silvia Snidero² and Dario Gregori³

¹Department of Clinical and Biological Sciences, University of Torino, Italy

²Department of Statistics and Applied Mathematics “Diego de Castro”, University of Torino, Italy

³Unit of Biostatistics, Epidemiology and Public Health, Department of Cardiac, Thoracic and Vascular Sciences, University of Padova, Italy

Corresponding author:

Dario Gregori,

Unit of Biostatistics, Epidemiology and Public Health,
Department of Cardiac, Thoracic and Vascular Sciences,
University of Padova,

via Loredan 18, 35121 Padova, Italy.

Email: dario.gregori@unipd.it

Abstract

Risk Assessment is the systematic study of decisions subject to uncertain consequences. An increasing interest has been focused on modeling techniques like Bayesian Networks since their capability of (1) combining in the probabilistic framework different type of evidence including both expert judgments and objective data; (2) overturning previous beliefs in the light of the new information being received and (3) making predictions even with incomplete data. In this work, we proposed a comparison among Bayesian Networks and other classical Quantitative Risk Assessment techniques such as Neural Networks, Classification Trees, Random Forests and Logistic Regression models. Hybrid approaches, combining both Classification Trees and Bayesian Networks, were also considered. Among Bayesian Networks, a clear distinction between purely data-driven approach and combination of expert knowledge with objective data is made. The aim of this paper consists in evaluating among this models which best can be applied, in the framework of Quantitative Risk Assessment, to assess the safety of children who are exposed to the risk of inhalation/insertion/aspiration of consumer products. The issue of preventing injuries in children is of paramount importance, in particular where product design is involved: quantifying the risk associated to product characteristics can be of great usefulness in addressing the product safety design regulation. Data of the European Registry of Foreign Bodies Injuries formed the starting evidence for risk assessment. Results showed that Bayesian Networks appeared to have both the ease of interpretability and accuracy in making prediction, even if simpler models like logistic regression still performed well.

Keywords

Bayesian Network, children, classification trees, foreign body injuries, quantitative risk assessment

Introduction

Quantitative risk assessment (QRA) is the systematic study of decisions subject to uncertain consequences by means of tools and techniques of probability theory and statistics.¹ One of the key features of QRA is its effort to look at whole systems and not isolated parts. Each possible adverse event is followed through to its consequences and at the same time the consequences of different adverse outcomes can be combined.

A wide range of techniques have been developed to address the risk assessment problem. They can roughly be classified as engineering, statistical or causal modeling techniques.² The engineering approach is based on the idea that risk objectively exists³ and risk analysis is a tool to express it by probabilities and expected values. Engineering techniques are mainly devoted to simulate the behavior of the system which is going to be assessed. In health risk assessment, Compartmental flow models and other continuous simulation models, Monte Carlo uncertainty models, Discrete-event simulation models are among the most common techniques, while Fault trees and Event trees are a major tool in safety and reliability analysis.⁴⁻⁷

Statistical risk modeling relies on observed data, on covariates and responses, rather than attempting to simulate the causal process that leads to the adverse outcome. The task is challenging because typically it is not known (1) which aspects of covariates are relevant to the response; (2) the mathematical form of the relation among variables and response probabilities; (3) how unobserved variables can affect the observed relation.² Among the techniques that have successfully been applied in risk analysis across various disciplines, there are logistic regression, artificial neural network and classification and regression trees.^{8,9}

While statistical and engineering approaches are complementary, causal modeling can combine elements of both. The risk assessment problem has been successfully addressed in a wide range of application domains using Bayesian Networks (BNs),^{10,11} which offer the benefit of explicitly model causal factors. The success of BNs in QRA is mainly due to their capability of: (1) combining in the probabilistic framework different types of evidence including both subjective beliefs and objective data, (2) overturning previous belief in the light of the new information received and (3) making predictions with incomplete data. The compositional modeling characteristic of the engineering approach is captured by the ingoing-outgoing relations in the network, while the conditional probability distribution of each variable may be determined by machine-learning algorithms.

In this paper, the comparison of different statistical modeling techniques developed in the framework of QRA is proposed. Unsafe consumer products are involved in

thousands of injuries in children, which can be caused for example by ingestion of batteries or broken plastic parts of toys and can have severe consequences.¹² The aim of this work consists thus in evaluating which modeling techniques, among those presented, can be better applied to assess the risk of such foreign bodies (FB) injuries in children. In general, probabilistic methods enable the characterization of the risk posed by products' characteristics such as their size or shape.¹³⁻¹⁵ In this paper, we focused on the application of statistical modeling techniques for the assessment of the risk to experience a severe injury. In particular, we were aimed at characterizing consumer and product's features and the surrounding circumstances that lead to hospital admissions.

In order to achieve this objective, data from the European Registry of Foreign Bodies Injuries "Susy Safe" was used. The "Susy Safe" Registry is a European Union funded registry, which has been established for the collection of FB injuries in children aged 0–14 and it is aimed at describing the clinical pattern and the public health burden of those injuries.

Following a short presentation of the data source, the implementation of the modelling techniques for QRA was described. Performance of the models (area under the ROC curve, sensitivity and specificity) were compared and a sensitivity analysis was carried out in order to

determine which explanatory variables have the most influence on the injury severity, thus giving an insight on how different modelling techniques may provide different explanation of the phenomenon under study. Finally, results were summarized and the benefit and limitations of each approach were discussed.

2. Materials and methods

2.1 Data source

The European Registry of Foreign Body Injuries "Susy Safe"¹⁶ collected data on FB injuries in children aged 0–14 according to the International Classification of Disease ICD9-CM 931-935. A total of 7296 cases were registered in 28 European hospitals at the end of March 2007. Data encompassed four main aspects of the FB injuries: (1) the characteristics of the children (age, gender); (2) the characteristics of the object (shape, rigidity and dimensions); (3) the circumstances of injury (presence of parents; activity performed by the child immediately before the accident); (4) hospitalization's details.

With regard to the FB dimensions, volume (calculated as the volume of the smallest regular geometrical solid containing the FB) and ellipticity (representing the

ratio between the maximum and the minimum size reported) were considered.

In order to analyze data in QRA's framework, children's age and gender were considered as control variables whereas FB characteristics (volume, ellipticity, shape and pliability) and the circumstances of the injury (adult presence or absence and the activity performed by the child) were considered as the key variables of interest, since they are key factors weighing on prevention strategies to avoid injuries or mitigate their severity.

2.2 Statistical methods

BNs and four other predictive models were implemented to quantify the risk of experiencing a severe injury. According to the report drafted from the e Consumer Affairs (CA) Directorate of the Department of Trade and Industry (DTI),¹⁷ an injury was defined "severe" when the injured child was hospitalized for at least 1 day.

After building the models, a 10-fold cross validation repeated 20 times was performed to evaluate their performance, which was summarized by the area under the ROC curve, the sensitivity and the specificity.^{18,19} Model fitting and model validation, with the exception of BNs, were carried out using R version 2.8.1.²⁰

Furthermore, for each model the mutual information (MI) was used to identify the variables that have the greatest influence on the risk of severe injury. MI allowed measuring the effect of each variable on the *Hospitalization* (yes/no) outcome variable and it was calculated as

$$I(hospitalization; Y) = \sum_{hospitalization_i \in \{0,1\}} \sum_{y \in Y} p(hospitalization_i, y) \log \frac{p(hospitalization_i, y)}{p(hospitalization_i)p(y)}$$

Where $p(hospitalization_i, y)$ is the joint distribution of the binary outcome and the variable Y and $p(hospitalizaion)$ and $p(y)$ their marginal probability distribution.^{21,22} Continuous variables were discretized. Age was binned into 5 bins and volume and ellipticity were binned into 10 bins. The number of bins was chosen in order to range between $1 + \log_2 n$ and $\sqrt{n}^{2.3}$. It has been shown that when the number of bins is within this range, the MI²⁴ outperforms other distance measures. Finally, the MI of each variable was normalized by the entropy of the outcome:

$$H(hospitalization) = - \sum_{i \in \{0,1\}} p(hospitalization_i) \log p(x_i)$$

2.2.1 Bayesian Networks

A BN is a graphical representation of the joint probability distributions over a set of random variables. It consists of a series of nodes representing variables connected by arrows forming a graph that has no cycles. The arcs specify the independent assumptions holding between random variables.

The resulting network is known as a directed acyclic graph (DAG).^{25,26} Each node in a BN is associated with a set of probability tables. For those nodes without ingoing arcs, the probability distribution is a prior distribution which requires supplying a set of initial values.

Different strategies can be adopted to build BNs²⁷: (1) both structure and probability tables can be generated from data; (2) both structure and probability tables are elicited from experts; (3) expert knowledge and objective frequency data can be combined, for example BN's structure can be defined by domain experts and probability tables can be learned from the data.

In this work, two BNs were implemented: (1) a first one (BN1) was completely generated from data; (2) a second one (BN2) was built using causal knowledge from otorhinolaryngologist physicians to model the structure^{28–30} and data to learn probabilities.

For BN1, the Greedy thick-thin algorithm³¹ was performed for learning the structure of the network. Since the software requires discrete variable, continuous variables were discretized on the basis of quintiles (*Age* and *Volume* variables) and tertiles (*Ellipticity* variable). Structure learning was carried out using GeNie.³²

The learning of probability tables and the validation phase were both carried out using Netica,³³ which allowed for specifying a fading factor in order to treat more recent cases with a higher weight than older ones.

For BN2, the implementation was entirely carried out using Netica.³³

2.2.2 Artificial neural networks

Artificial Neural Networks denote a set of information processing paradigm inspired on the biological nervous system behaviour. The multilayer perceptron (MLP) is the most popular neural architecture where neurons are grouped in layers and only forward connections exist.³⁴

Several feed-forward neural networks architectures with back-propagation learning method were implemented.³⁴ All neural networks contained from 10 to 25 neurons in a single layer and one neuron in the output layer. In all calculations, the layers were fully connected. The least number of misclassifications given as the average on the validation datasets was obtained for the network with 17 neurons in the hidden layer.

2.2.3 Classification trees

Classification tree (CT) is a nonparametric method based on recursive partitioning of a sample into subgroups. At each step the most significant predictor is used to split the sample into subset until no improvement is achieved in the classification accuracy.³⁵

Two classification trees have been implemented. A first one (CT1) was developed by the standard binary recursive partitioning using all predictor variables described in Table 1.³⁶

Table 1. Definition of variables and their states in the Bayesian network. Continuous variables were discretized on the basis of quintile (Age and Volume nodes) and tertile (Ellipticity node).

Node description	Variable description	State description
------------------	----------------------	-------------------

Age	Continuous	Age class: 0–1, 2, 3–4, 5–6, 6–14
Gender	Discrete	Female, male
Location	Discrete	ICD931-935
Hospitalization	Discrete	No, Yes
Complications	Discrete	No, Yes
Extraction technique	Discrete	Aspiration, bronchoscopy, endoscopy, operation, microotoscopy, otoscopy, other
Foreign body type	Discrete	Accessorize, arthropod, battery, bean and pea, bone, button, capo, coins; cotton, earplug, fruit and stone; jewelers; metal; nut; other inorganic; other organics; papers; pearl, ball and marble; pebble; pins and needle; plastic; stationery; stick, sweet, toys
Shape	Discrete	2D circle; 3D; cylinder; needle shape; other; spherical
Rigidity	Discrete	Pliable, stiff, semi-stiff
Ellipticity	Continuous	1, score from 1 to 2, greater than 2
y	Continuous	Score up to 33.5, score from 33.5 to 65.4, score from 65.4 to 140, score from 140 to 400, up to 4710
Adult Presence	Discrete	Adult absent, Adult present
Activity before accident	Discrete	Eating, playing, other

To avoid over-fitting, 10-fold cross validation was used to determine the optimal size of the tree. The best size was selected according to the 1SE rule, by which the largest tree with cross-validated error within one standard deviation of the minimum was chosen.³⁷

A second classification tree (CT2) was implemented by means of a hybrid approach, exploiting a BN. Following Cox's establishments,³⁸ a BN was implemented and sensitivity analysis was carried out. All variables whose MI had the outcome node (Hospitalization) fully explained by other variables or inconsistent with the hypothesis of causality were discarded. Finally a classification tree considering the outcome node and its minimal set of predictors, consisting in outcome node's parents, its children and children parents, was implemented. These pre-processing data procedures allowed to eliminate variables statistically associated with the outcome variable only due to confounding. In fact, in a causal graph $X \leftarrow Z \rightarrow Y$, the parent node Z is a confounder since it explains away an apparent association between X and Y . Thus, including outcome nodes' parents in the classification tree permitted to avoid these situations.³⁸

2.2. 4 Logistic regression

Logistic regression (LR) model is a statistical tool widely used to fit probability of an event by a linear function of the explanatory variables.³⁹

A logistic regression model was constructed using backwards variable elimination at a significant level of 0.05. The backward variable elimination was based on sequential elimination of variables from an initial model consisting of all the predictor variables. At each step the variable which resulted in the greatest reduction of the AIC criterion was removed from the model. The rule of eliminating variables followed on until no further significant reduction of the AIC was obtained.¹⁹ Interaction among variables was checked in a similar way.

2.5 Random Forest

Random forest (RF) is a collection of classification trees.⁴⁰ Each tree is grown using a bootstrap sample from the original data. About two-thirds of the data are used to construct the classification tree, whereas the remaining Out-Of-Bag (OOB) data, which is left out, is used to obtain unbiased estimates of correct classification rates and variable importance.

Bootstrapping procedures are carried out for building an ensemble of trees with a reduced dependence among them. When building a classification tree, for each node of the tree, the RF algorithm selects some variables (the number of variables to select is usually taken to be the square root of the total number of variables) and uses only them to determine the best possible split at a single node – which is determined by the independent variable that best divides the sample in that node into two subgroups, each with the most pure membership using the Gini index as the splitting criterion.

Each tree in the forest is grown using the bootstrapped sample, and the OOB part is thus processed by the grown tree. This gives rise to classification for each point in the OOB part of that bootstrapped sample, meaning that about one third of the trees in the random forest give a prediction for each point in the original data. The final classification of a particular data point is decided on the basis of majority vote. The unbiased estimates of true classification rates are calculated by comparing the OOB set classification made by the forest to the experimentally observed classes to which the data points belong. The random forest package in R was used to implement the model.

3 Results

3.1 Input variables

In Table 1, input variables provided by the Susy Safe registry were listed in the “*Node description*” column. Variables *Age* and *Gender* recorded demographic characteristics of the injured child; *Location* reported the location of the foreign body, which caused the accident using International Classification of Disease (ICD-9) codes: ICD931 (FB in the ears), ICD932 (FB in the nose); ICD933 (FB in the pharynx and larynx); ICD934 (FB in the trachea, bronchi and lung); ICD935 (FB in the mouth).

The details of hospitalization were provided in two variables: (1) *Hospitalization*, which recorded whether the child experienced at least 1 day of hospitalization; (2) *Complications*, which recorded whether the child experienced complications, e.g. obstructions, pneumonia esophageal atresia, nasal odorous discharge.

FBs were described by shape, rigidity and size, according to Rimell’s definition.⁴¹ With regard to

their shape, FBs were assigned to one of the following categories: 2D circle (e.g. some pieces of paper), 3D (e.g., pen cap), cylinder (e.g. coins), needle shape (e.g. pins and needles), spherical (e.g. balls and pebble) and other shapes.

With regard to the size, when the dimensions of the object (given in mms) were reported, the volume was calculated as follows: for 3D objects the volume of a parallelepiped was calculated considering the length of the axes; for spherical objects the volume of a sphere was calculated by the diameter reported; for 2D circle objects, the volume was approximated by that one of a cylinder with height 1 mm. Such volume measures represent how much space takes up the smallest geometrical figure containing the irregular-shaped foreign body. In addition, for three-dimensional not spherical objects, the ellipticity, i.e. the ratio between the longest axis and the shortest axis, was computed.

In Table 1, modalities of the discrete variables were reported (column 3), along with the classes used to discretize the continuous variables.

The structure of the BNs BN1 and BN2 is depicted in Figures 1 and 2, respectively (see Table 1 for the description of nodes); classification trees CT1 and CT2 are shown in Figures 3 and 4,

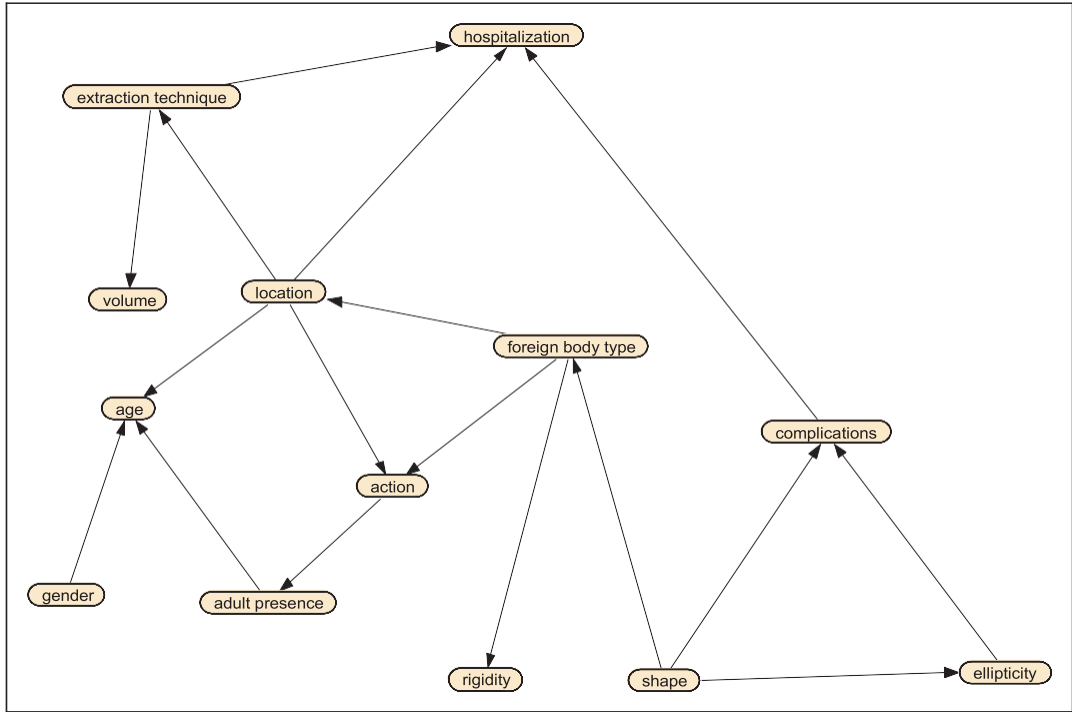


Figure 1. Structure of Bayesian network (BN) completely generated from data (BN1). Labeled rectangles are node (model variables); arrows represent conditional dependence relationships. Since BN1 is completely generated from data, the arrows between two nodes do not imply causation but just the existence of a relationship, which is given by a conditional probability table.

respectively. In implementing the classification tree CT2, potential confounding effects with other variables were eliminated by means of data pre-processing described in section 2.2.3. Thus, in this case, the relationship among severe injuries and other variables was potentially interpretable as causal. Logistic regression parameter estimates are shown in Table 2.

In Table 3, the performance of each of the models is summarized with the Area under the ROC curve, sensitivity and specificity along with their 95% confidence intervals. The ranking of features based on the MI computed among *Hospitalization* and all other model variables is shown in Table 4, giving thus a measure of the relative importance of the variables as predictor of severe injuries.

It could be observed that in the BNs, nodes which are closer to the *Hospitalization* node (Figures 1 and 2) showed a greater impact in predicting injury severity. Conversely, the influence of nodes farther away tended to be diluted due to the uncertainty introduced by the intermediate nodes, in a phenomenon already known

and described in the literature.²¹ For the logistic regression model, backward variable elimination yielded a reduced model with 6 out of the 12 original variables.

3.2 Scenarios definition and prediction

The combination of events, features and processes causing diverse natural phenomena could be taken as a scenario. The capability to predict scenarios and compute an occurrence probability is a

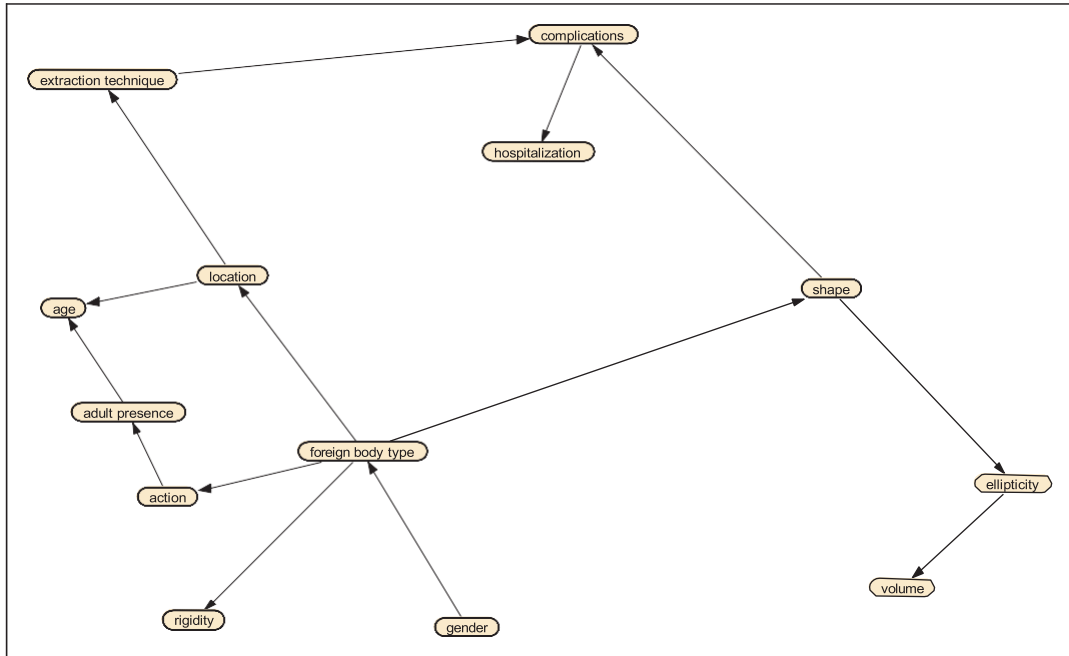


Figure 2. Bayesian network (BN) modeled using causal knowledge of the phenomenon derived from otorhinolaryngologist (BN2). Labeled rectangles are node (model variables); arrows represent conditional dependence relationships.

valuable tool for risk assessment because it allows for extrapolation of hazard and prevention. BNs can handle this feature in a very straight-forward way. To illustrate this point we calculated the risk of hospitalization of a few scenarios that may be encountered in the clinical practice.

In Figure 5, an example about how BNs deal with scenarios is presented. In this example, BN1 has been considered since it is the BN model that achieves the best accuracy (Table 3). After setting the evidence (a male who had an injury while he was playing with a spherical shaped object) the probability of being hospitalized was computed making use of the Bayes Theorem. Thus, given the injury occurred, the probability to be hospitalized is 83.3%. Entering new evidence made it possible to

update probabilities. For example, if a needle-shaped foreign body, e.g. a fishbone, is swallowed while a child is eating, the probability of an injured male to be hospitalized was about 40%, whereas for an injured female was slightly lower (37%).

Finally, since the BN completely generated from data (BN1) outperformed all other models (Table 3), it was used to construct a set of risk profiles (Table 5). In fact, given the type of the foreign body, its shape, rigidity and volume, child's age and gender, the probability of observing an injury was computed. Also the most probable location of the foreign body along with the most probable extraction techniques required were reported in addition to the probability of experiencing a hospitalization.

4 Conclusion

In this paper, a comparison among techniques widely used in QRA was proposed. An approach based on BNs was adopted to carry out a risk analysis on foreign body injuries in children. BNs

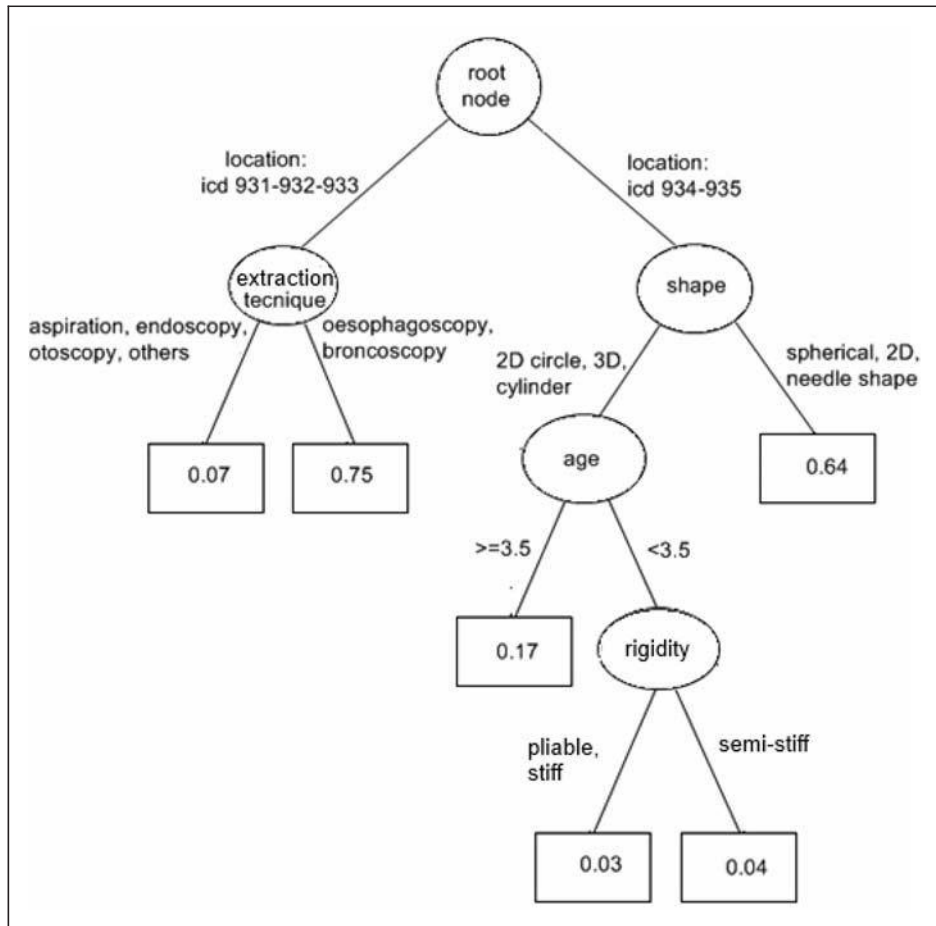


Figure 3. Classification tree (CT1). Prior probabilities at each group have been treated as equal. Terminal nodes are symbolized by rectangles; non-terminal nodes, by ovals. Splitting criteria are specified in the nodes. The probability of experiencing a hospitalization is given within the node according to the set of rules specified.

performance was compared to a set of competing statistical risk modeling methods: (1) logistic regression models (LR); (2) artificial neural networks, ANN; (3) classification trees, CT1 and (4) random forest. A hybrid approach using BNs in building classification trees (CT2) was also considered. Children's hospitalization was identified as the outcome measure of injury severity and it was studied in relation to the child's age and gender and accident details.

Artificial neural networks (ANNs) along with classification trees and random forest are a rich tool in dealing with noisy or incomplete data. A drawback of ANNs is, however, that there are no standard methods for constructing the architecture. In this study, we set up a single hidden layer feed-forward neural network, which is the

common type encountered in the literature.³⁴

The classification tree approach is also a method extremely robust to the presence of irrelevant variables and variables with little predictive value. Besides a non-impressive total accuracy, standard classification tree (CT1) showed to be capable of better identifying injuries at a higher risk of severity (96% of sensitivity). Furthermore, contrary to ANNs, which are “black box” models of difficult

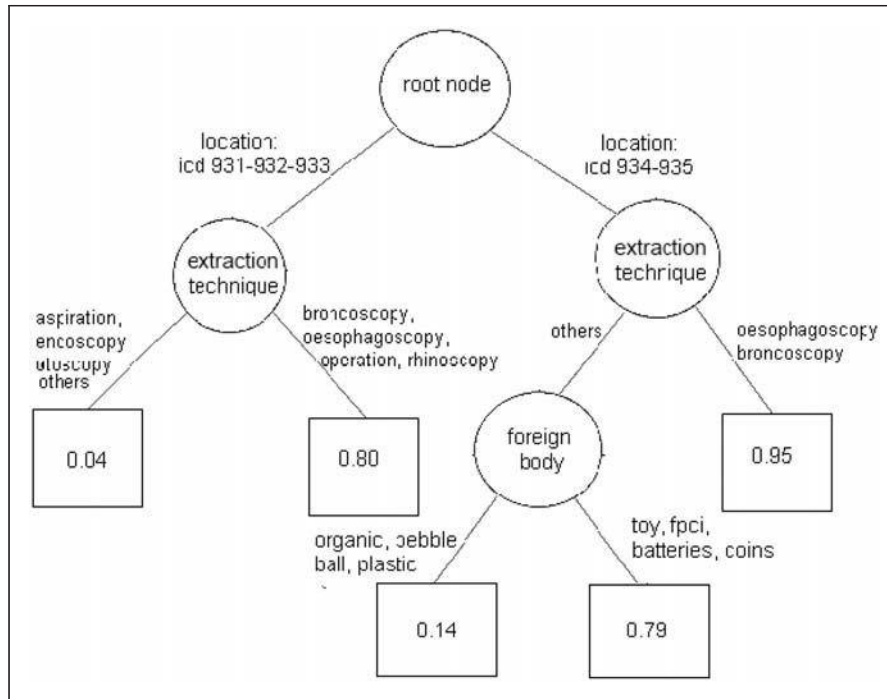


Figure 4. Classification tree built using a hybrid approach (CT2). A Bayesian network was completely generated from data. Then a standard classification tree was fit between the outcome (hospitalization) and the minimal set of predictors given by the parents in the Bayesian network.

interpretation, it provided a way to extrapolate decision rules for achieving threat reduction in the form of a ready-understandable flowchart.

Like ANNs, random forests do not present a ready-understandable flow-chart. They are an ensemble method, which reduces variability of trees by averaging multiple trees. Thus, as expected, it showed a slightly better specificity and a lower sensitivity than CTs, achieving an overall performance which made it comparable to ANNs.

Logistic regression is a popular statistical method, owing largely to its simplicity

and the interpretability of the estimated parameters, which can generate excellent prognostic models. Although LR is not adept at modeling grossly nonlinear complex interaction, in our study it showed indeed its ability to capture non-linear effects outperforming ANNs and standard CT, which also were affected by a low specificity.

Two different strategies were chosen to implement BNs: (1) a BN in which both the structure and probability tables were generated purely from the data (BN1); (2) a BN in which the structure was defined by experts (mainly otorhinolaryngologists) of the phenomenon, whereas the probability tables were entirely learned from the data (BN2). One of the main differences that arose between the two models was the role of the *Extraction technique* and *Complications* variables. In a causal representation of the phenomenon, extraction methods can be potential cause of complications, which is the most important predictor of injury severity. This fact did not arise in the BN completely generated from data (BN1) as well as in the automatically generated models. Indeed, inspecting the

Table 2. Results of logistic regression model.

Variable	Effect	p-Value
<i>Location</i>		
ICD933	1.35	0.029
ICD934	3.79	<0.001
ICD935	2.44	<0.001
<i>Gender</i>		
Female	0.89	0.045
<i>Extraction technique</i>		
Aspiration	0.37	0.053
Bronchoscopy	3.35	<0.001
Oesophagoscopy	3.43	<0.001
Operation	2.64	0.023
Otoscopy	0.03	0.042
<i>Shape</i>		
2D circle	1.36	<0.001
Cylinder	1.45	0.003
Needle shape	1.47	<0.001
Other	1.82	<0.001
<i>Rigidity</i>		
Pliable	0.87	0.04
Semi-stiff	0.80	<0.001
Age	0.80	0.003

Table 3. Area under the ROC curve was used to assess the performance of the models: Bayesian Networks (BN1 and BN2); logistic regression (LR), neural networks (ANN); classification trees (CT1 and CT2); random forest (RF).

	AUC	SENS	SPEC
BN1	92.31% (89.94-94.68)	95.19% (93.1-97.28)	90.06% (88.08-92.04)
CT2	91.5% (89.39; 93.61)	87% (85.03; 88.97)	32% (29.82; 34.18)
LR	87.03% (84.99-89.05)	89.2% (86.5-91.9)	83.1% (80.75-85.45)
ANN	74.45% (71.25-77-65)	81.42% (77.52-85.32)	55.16% (51.51-58.8)
CT	72.29% (70.27-74.31)	96.14% (94.06-98.22)	41.74% (39.58-43.9)
BN2	71.59 % (68.48-74.7)	91.12 % (87.9-94.44)	40.17 (36.96-43.38)
RF	86.11% (82.90-89.31)	73.52% (64.97-82.08)	57.15% (54.11-60.19)

registry, it has been recognized that many records reported symptoms related to the presence of a FB, such as pain, epistaxis or hearing loss, instead of clinical complications, and this indeed explained why extraction technique and not complications was the most influential variable in predicting severe injuries.

Our analysis has shown that BN1 and the hybrid approach CT2 outperformed all methods in terms of accuracy. Opposing to CT2, the advantage of BNs relied on the fact that complex relationships among factors were explicated in a graphical model, which incorporated uncertainty via the conditional probability associated to each node.²⁶ Indeed, BNs gave a picture of the influence of critical factors on the injury severity. Results from ranking of variables suggested the conclusion

Table 4. Sensitivity analysis was carried out on hospitalization variable to determine the covariates that have the most influence on the injury severity.

BN1		CT2		BN2		ANN		CT1		LR		RF	
Variable	%	Variable	%	Variable	%	Variable	%	Variable	%	Variable	%	Variable	%
Location	45.3	Location	68.6	Complications	75.1	Location	45.4	Location	37.2	Location	58.7	Location	50
Extraction technique	40.1	Extraction technique	52.7	Extraction technique	44.8	Extraction technique	25.5	Extraction technique	20.8	Extraction technique	36.4	Extraction technique	47.2
Foreign body type	17	Foreign body type	43.8	Shape	43.5	Shape	12.4	Shape	13.2	Shape	10.5	Volume	20.1
Shape	11.3			Foreign body type	38.2	Foreign body type	6.8	Age	5.9	Age	7.4	Age	11.8
Action	2.94			Location	37.7	Rigidity	3.7	Rigidity	3.1	Rigidity	4.3	Shape	4.1
Consistency	2.62			Ellipticity	31.4	Volume	2.3	Foreign body type	2.4	Gender	3.2	Foreign body	3.8
Adult presence	0.059			Action	27.6	Action	1.09	Action	1.2			Action	1.7
Ellipticity	0.0245			Rigidity	27.6	Age	1.2	Volume	0.98			Adult presence	1.1
Age	0.021			Volume	24.2	Ellipticity	0.4	Ellipticity	0.3			Ellipticity	1.0
Volume	0.01			Adult presence	17.8	Adult presence	0.5	Adult presence	0.19			Rigidity	0.9
Gender	0.002			Age	16.5	Gender	0.09	Gender	0.03			Gender	0.7
Complications	0.009			Gender	2.13	Complications	0.002	Complications	0.001			Complications	0.7

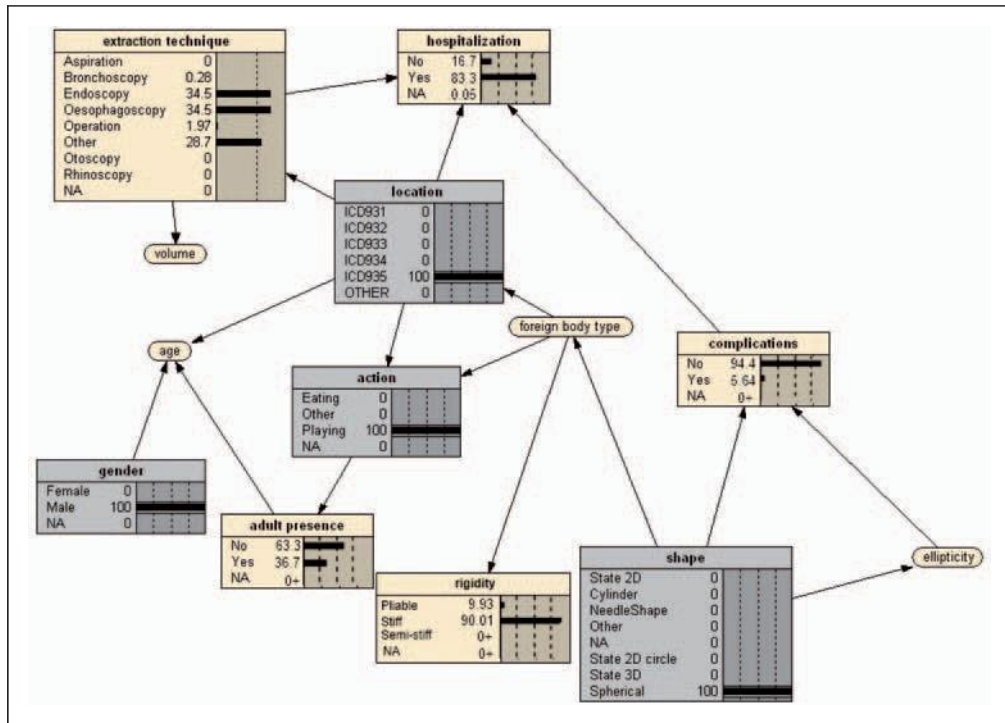


Figure 5. The scenario of a male who had an injury while he was playing with an object with spherical shape is depicted using BN1. Foreign body (FB) was located in the mouth, esophagus or stomach (ICD935). Hospitalization node returned the probability the injury lead to hospital admission (83.3%). Probability table associated to *Foreign body type* node (not showed to allow for a better visualization) gives that there is 89% of probability the child is playing with a button. Also, there is 63.3% of probability the injury occurred in absence of an adult (see *adult presence* node).

that models do not always give the same interpretation for the same covariates, according to other studies,^{42,43} rather they provide a different framework for the explanation of the evidence.

Thus, a tool which allows for a ready interpretation of relationships among risk factors is of great use. In the framework of risk analysis, this capability makes BN a competing alternative to other approach, such as logistic regressions, which are often preferred since they provide easily interpretable results. The possibility to analyze different scenarios is an import feature which allows to asses the effect of FB's characteristics and incident's circumstances, such as adult absence/presence, on the risk of experiencing a severe injury, once the injury has occurred. In fact, probability updating after setting evidence can be used to identify in a straightforward manner the characteristics of unsafe products or the importance of adult surveillance in

mitigating injury severity.

However, it should be noted that the logistic regression model showed an overall classification accuracy which was not far behind BNs, confirming to be an efficient model for classification task, despite its simplicity.

Even if BNs can be considered for causal modeling, in this study they were considered for association analysis only. The absence of an independent sample to externally validate the models constituted the major limitation of this study.

Table 5. Predicted probability of observing evidence on foreign body and children characteristics based on BN1. The probabilities of the most probable FB location (ICD) and extraction technique required are reported along with the probability of experiencing a hospitalization given that an injury occurred.

Observation pattern								Probability of observing evidence	The most probable ICD	The most probable extraction technique	Probability of hospitalization
N	Foreign body type	Age	Gender	Shape	Rigidity	Volume	Ellipticity				
1	Batteries	2	m	spherical	Stiff	70	1	6.7	935 (84%)	Endoscopy (78%)	88%
2	Pebble	1	m	3D	Stiff	140	1.4	4.3	932 (67%)	Endoscopy (49%)	83%
3	Plastic	3	F	3D	Pliable	95	>2	7.3	932 (73%)	Endoscopy (33%)	27%
4	Fish bone	5	m	Needle shape	Pliable	140	>2	7.4	934 (61%)	Endoscopy (52%)	62%
5	Pearls	6	F	Spherical	Stiff	>400	1	10.1	934 (49%)	endoscopy (42%)	71%
6	Stationery	4	m	Cylinder	Pliable	33.5	>2	4.56	934 (62%)	Endoscopy (37%)	56%
7	Toy	2	F	Spherical	Stiff	102	1.3	3.8	933 (65%)	Other (31%)	53%
8	Nut	6	m	Spherical	Stiff	200	1	7.2	934 (89%)	Bronchoscopy (58%)	87%
9	Button	3	m	2D circle	Stiff	>40	2	2.4	935 (61%)	Other (66%)	39%
10	Stick	4	m	Needle shape	Stiff	NA	NA	1.3	933 (90%)	Other (71%)	18%

BN: Bayesian Network; FB: foreign body; ICD: International Classification of Disease.

4.1 Final remarks

While logistic regression models were found to be a simpler model, which still perform well comparably with other more complex statistical techniques, BNs, beyond outperforming all other models, offered some advantages in the context of QRA.

Since the “Susy Safe” surveillance registry is set up to constantly receiving new cases, we chose to treat BN as an adaptive net giving a higher weight to more recent cases with respect to the older ones. As a result, we built a BN that while receiving cases and updating information on foreign body features (size, shape and rigidity) was able to quickly respond to instances of changing product safety design regulation.

The capability of identifying relationships among variables is a key feature of BNs. In this analysis, the BN confirmed the role of FB’s type along with its shape and rigidity in determining the risk of severe injury, beside its location and the extraction procedure chosen by physicians. Moreover, the complex relationships among risk factors showed that there was not a single cause related to the severity of the injury but a more complex pattern of events contributing to the adverse outcome.

The conclusions drawn from this single comparative study are certainly not definitive.

References

1. The Royal Society. *Risk: analysis, perception and management: Report of a Royal Society Study Group*. London, 1992.
2. Cox LA. *Risk analysis foundations, models and methods*, ed. I.S.i.O.R.M. Science. Boston: Kluwer Academic Publisher, 2002.
3. Aven T and Kristensen V. Perspectives on risk: review and discussion of the basis for establishing a unified and holistic approach. *Reliability Eng Syst Saf* 2005; 90(1): 1–14.
4. Vose JD. The application of quantitative risk assessment to microbial food safety. *J Food Protect* 1998; 61(5): 640–648.
5. Wreathall J and Nemeth C. Assessing risk: the role of probabilistic risk assessment (PRA) in patient safety improvement. *Qual Saf Health Care* 2004; 13(3): 206–212.
6. Roy PK, Bhatt A and Rajagopal C. Quantitative risk assessment for accidental release of titanium tetrachloride in a titanium sponge production plant. *J Hazard Mater* 2003; 102(2-3): 167–186.
7. Clark LH, Setzer RW and Barton HA. Framework for evaluation of physiologically-based pharmacokinetic models for use in safety or risk assessment. *Risk analysis* 2004; 24(6): 1697–1717.
8. Floerl O, Inglis GJ and Hayden BJ. A risk-based predictive tool to prevent accidental introductions of nonindigenous marine species. *Environ Manage* 2005; 35(6): 765–778.
9. Levy JI, Wolff SK and Evans JS. A regression-based approach for estimating primary and secondary particulate matter intake fractions. *Risk Analysis* 2002; 22: 895–904.
10. Fenton NE, Krause P and Neil M. Software Measurement: Uncertainty and Causal Modelling. *IEEE Software* 2002; 10(4): 116–122.
11. Heckerman D, Mamdani A and Wellman M. Real-world applications of Bayesian networks. *Comm ACM* 1995; 38(3): 25–26.
12. Gregori D, et al. *Foreign body causing and prolonging hospitalization in children 0-14 age: early results from the ESFBI Study*. In: *Healthcare System and Patient Safety*. Florence, Italy: Taylor & Francis, 2005. Rider G, et al. Quantitative risk analysis. *Injury Contr Saf Promotion* 2000; 7(2): 115–133.
13. Rimell FL, et al. Characteristics of objects that cause choking in children. *J Am Med Assoc* 1996; 274(22): 1763–1766.
14. Stool D, Rider G and Welling JR. Human factors project: development of computer models of anatomy as an aid to risk management. *Int J Pediatr Otorhinolaryngol* 1998; 43(3): 217–227.
15. Gregori D. The Susy Safe Project A web-based registry of foreign bodies injuries in children. *Int J Pediatr Otorhinolaryngol* 2006; 70(9): 1663–1664.
16. DTI. *Choking risk to children under four from toys and other objects*. London: DTI, 1999.
17. Kohavi R. A study of cross-validation and bootstrap for accuracy

- estimation and model selection. *Proceedings of Fourteenth International Joint Conference on Artificial Intelligence (IJCAI)*, 1995. Montreal, Quebec, Canada: Morgan Kaufman Publishing, pp. 1137–1143.
18. Harrell JFE. *Regression modeling strategies*. New York: Springer, 2001.
 19. R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012.
 20. Cover TM and Thomas JA. *Elements of information theory*. xxiii, 2nd edn. Hoboken, NJ: Wiley-Interscience, 2006, p.748.
 21. Pearl J. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. San Mateo, California: Morgan Kaufmann Publisher, 1991.
 22. Law AM. McGraw-Hill series in industrial engineering and management science. *Simulation modeling and analysis*. xix, 4th edn. Boston; London: McGraw-Hill, 2007, p.768.
 23. Priness I, Maimon O and Ben-Gal I. Evaluation of gene- expression clustering via mutual information distance measure. *BMC Bioinformatics* 2007; 8: 111.
 24. Jensen FV. *Bayesian networks and decision graphs*. Springer Verlag, 2001.
 25. Pearl J (ed.) *Causality: models, reasoning and inference*. Cambridge: CU Press, 2000.

26. Joseph A, Fenton NE and Neil M. Predicting football results using Bayesian nets and other machine learning techniques. *Know-Based Syst* 2006; 19(7): 544–553.
27. Matheny AP. In: Routh D (ed.) *Accidental injuries*, in *Handbook of pediatric psychology*. New York: Guilford, 1988.
28. Rivara FR and Mueller BA. *The epidemiology and causes of childhood injuries*. *J Soc Issues* 1987; 43(2): 13–21.
29. Health BMo. *A report on the health of British Columbians. Provincial health officer's annual report. Feature report: The health and well-being of British Columbia's children*. B.M.o. Health, Editor: Victoria, BC, 1997.
30. Hackerman D. A tutorial on learning with Bayesian networks. In: Jordan MI (ed.) *Learning Graphical models*. Dordrecht, Netherlands: Kluwer, 1998, pp. 301–354.
31. Decision Systems Laboratory – University of Pittsburgh. *GeNIe 2.0*, <http://www.sis.pitt.edu/genie/> (2006, assessed 3 September 2012).
32. Norsys Software Corporation. Netica v3.18, <http://www.norsys.com/netica.html> (2006, assessed 3 September 2012).
33. Ripley BD. *Pattern recognition and neural networks*. xi, Cambridge: Cambridge University Press, 1996, p.403.
34. Breiman L, et al. *Classification and regression trees*. New York: Chapman and Hall, 1984.

35. Berchialla P, et al. Predicting severity of foreign body injuries in children in upper airways: an approach based on regression trees. *Risk Anal* 2007; 27(5): 1255–1263.
36. Venables WN and Ripley BD. *Modern applied statistics with S*, 4th edn. New York: Springer, 2002.
37. Cox LA. Detecting causal nonlinear exposure-response relations in epidemiological data. *Dose Response* 2006; 4(2): 119–132.
38. McCullagh P and Nelder JA. *Generalized linear models*. 2nd edn. Monographs on statistics and applied probability; 37. London: Chapman and Hall. xix, 1989, pp. 511.
39. Breiman L. Random forests. *Mach Learning* 2001; 45(1): 5–32.
40. Rimell FL, et al. Characteristics of objects that cause choking in children. *JAMA* 1995; 274(22): 1763–1766.
41. Gregori D, et al. Using data mining techniques in monitoring diabetes care. *The simpler the better?* J Med Syst 2011; 35(2): 277–281.
42. Bartfay E, Mackillop WJ and Pater JL. Comparing the predictive value of neural network models to logistic regression models on the risk of death for small-cell lung cancer patients. *Eur J Cancer Care (Engl)* 2006; 15(2): 115–124.